Financ Mark Portf Manag (2012) 26:247–267
DOI 10.1007/s11408-012-0183-4

The pricing of idiosyncratic risk: evidence from the implied volatility distribution

Stephan Süss

Published online: 14 April 2012

Abstract A recent strand in the literature has investigated the relationship between idiosyncratic risk and future stock returns. Although several authors have found significant predictive power of idiosyncratic volatility, the magnitude and direction of the dependence is still being debated. Using a sample of all S&P 100 constituents, we identify positive risk premia for option-implied idiosyncratic risk. Depending on the model used to identify unsystematic risk, we observe a statistically and economically significant average annual premium of 1.72 percent. To investigate whether this impact is driven by the definition of idiosyncratic risk, we extend the pricing kernel by implied skewness. Using a double-sorting procedure, we show that the compensation of unsystematic risk is mainly driven by firms with high positive implied skewness.

Keywords Idiosyncratic risk · Implied volatility · Implied skewness · Principal portfolios · Random matrix theory

JEL Classification G12

1 Introduction

Research on the impact and direction of risk on the cross-section of future asset returns has a long history in the financial literature. Traditional asset pricing theory predicts that only systematic risk (as defined by the specific form of the pricing kernel) should have an impact on returns, as unsystematic or idiosyncratic risk can be diversified by the investor. Empirical tests of this hypothesis frequently fail to confirm this perception of the relation between risk and return. Early studies, such as Lintner

S. Süss (✉)
Swiss Institute of Banking and Finance, University of St. Gallen, Rosenbergstrasse 52, 9000 St. Gallen, Switzerland
e-mail: stephan.suess@unisg.ch

(1965) and Lehmann (1990), find a positive relationship between the residuals of the parsimonious Market Model and realized asset returns.[1] However, the famous studies of Ang et al. (2006, 2009) find a significant and very robust negative impact of unsystematic risk on asset returns. Thus, the nature and direction of the relationship between idiosyncratic risk and returns is still under debate. In the past, research has taken several different routes to address this issue.

Merton (1987) considers a simple model economy with incomplete information to predict a positive impact of idiosyncratic risk on asset returns. A similar approach is suggested by Barberis et al. (2001), who model the loss aversion of a representative investor depending on the past total investment performance. Their model predicts a positive pricing of idiosyncratic risk. Both approaches rely on investors holding under-diversified portfolios, which has been confirmed empirically by, for example, Goetzmann and Kumar (2008). In a sample of more than 40,000 equity investment accounts, the median investor has no more than three different stocks in her portfolio.

A different route taken in the literature to address the problem of mis-specification of the classic factor models is to investigate the implications of higher-order return distribution moments, such as skewness and kurtosis, for capital asset pricing.[2] However, the number of return-generating factors, as well as the magnitude and direction of their impact, is rarely known ex ante. Second, estimation of the expected correlation between asset returns and the identified factors remains a highly challenging task, especially in the presence of correlation risk premia as found in Driessen et al. (2009).

This study uses Random Matrix Theory (RMT) to identify systematic and unsystematic implied equity return risk. Our method is able to handle the problem of correlated risk components. Furthermore, this method can cope with time-varying numbers of risk sources and guarantees positive semi-definiteness of the resulting covariance matrices for systematic and idiosyncratic risk components. We apply our method to the constituents of the S&P 100 and find significantly positive idiosyncratic risk premia for the January 1996 to October 2010 period. With an average annual premium of 1.72 percent, it is statistically and economically significant. Extending our pricing kernel by implied skewness, we find that the premia on idiosyncratic risk can to a large extent be explained by higher moments of the implied return distribution.

Our article is organized as follows. Section 2 introduces our volatility and correlation estimators. Section 3 outlines details of our dataset. Empirical results are given in Sect. 4. Their robustness is investigated in Sect. 5. Section 6 concludes.

2 Methodology

2.1 The estimation of the return covariance matrix

The identification of unsystematic risk premia is highly dependent on the definition of systematic return factors, that is, the definition of the pricing kernel. To measure risk,

[1] For more recent studies see, among others, Fu (2009), and Diavatopoulos et al. (2008).

[2] See, e.g., Vanden (2006) and the references therein.

we need the best possible predictor for the future return covariance matrix. Here, we have different choices for the estimation of the return volatility vector, as well as different estimators for the correlation matrix. Roughly speaking, volatility estimators can be grouped into estimators based on historical time-series observations or implied volatility estimators. Reliance on risk-neutral estimates requires careful consideration of several points. At least two opposing effects come into play. On the one hand, implied estimators are based on empirically observable market prices. Therefore, they may be considered purely forward-looking variables with a high ability to reflect changing market conditions in a timely way. Thus, implied return moments are less prone to the statistical inertia of sample return time series. On the other hand, estimators under the risk-neutral measure reflect investor sentiment at the time of portfolio construction. As such, they can substantially differ from the realized values in the future. It is well-known that implied volatility typically overestimates future realizations. This phenomenon entered the financial literature as volatility risk.[3] In addition, liquidity effects contained in option prices may distort implied moment estimators. To investigate the relative importance of the two opposing effects, we implement estimators on the historical, as well as the implied probability measures. Summarizing, it is not clear ex ante whether estimators relying on implied volatility outperform those relying on the sample time series.[4]

The remainder of this section outlines our correlation matrix estimators. Our study uses three different methods. First, we calculate the correlation matrix based on sample time series of 250 calendar days ($\rho^{\mathbb{P}}$). Note that the length of the estimation window must exceed the number of sample firms. Otherwise, our matrix estimator would be singular. However, we are aware that our estimator is prone to data inertia.

A promising alternative to the use of the sample time-series estimator is the HETIC model of Buss and Vilkov (2011) (ρ^{HET}). They try to infer information from the option market by estimating a single-factor model for the correlation matrix. The idea is simple: the return variance of a portfolio or index M consisting of N assets with weights w_i, $i \in \{1, \ldots, N\}$ can be written as

$$\sigma_M^2 = \sum_{i=1}^{N} w_i^2 \sigma_i^2 + \sum_{i=1}^{N} \sum_{j \neq i}^{N} w_i w_j \sigma_i \sigma_j \rho_{ij}. \tag{1}$$

This equation holds true under the physical, as well as the risk-neutral measure.

The identification of the implied correlation matrix requires the estimation of $N(N-1)/2$ parameters. Uniqueness of a risk-neutral covariance matrix estimator

[3] See, e.g., Bakshi and Madan (2006), Bollerslev et al. (2009), and the references therein on volatility and variance spreads. This collection is by no means exhaustive. Furthermore, see Driessen et al. (2009), Krishnan et al. (2009), and Buss and Vilkov (2011) for empirically observable correlation risk premia.

[4] A great deal of literature compares the predictive power of implied and historical measures for volatility. Early research in this field found that estimators based on implied volatility typically outperform estimates based on historical volatility. See, e.g., Latané and Rendleman (1976), Schmalensee and Trippi (1978), and Beckers (1980). Contrary to these findings, Canina and Figlewski (1993), Day and Lewis (1992), and Lamoureux and Lastrapes (1993) deny predictive power of implied volatility. More recent work, for example, includes Blair et al. (2001), Lehar et al. (2001), Jiang and Tian (2005), and DeMiguel et al. (2011). Buss and Vilkov (2011), among others, find a high predictive power of intraday return data to estimate volatility. Others find a reasonable empirical fit for component models, such as Zhu (2009).

would require observing option values for each possible asset pair. As noted by Buss and Vilkov (2011), implied correlation estimation has been a challenging task in the literature. The high dimensionality of the estimation problem requires further restrictions to make the implied correlation matrix identifiable. Driessen et al. (2009) use equal correlations for each off-diagonal element to study the correlation risk premium. This method is inappropriate for our task as it does not identify the diversification potential of single assets. A more appropriate method has been suggested by Buss and Vilkov (2011). They use an affine single-factor setting for modeling correlation risk premia. This allows preserving the heterogeneity of the correlation matrix. The time-t correlation risk premium $CRP_{ij,t}$ between two asset returns is defined as the difference between the objective ($\rho^{\mathbb{P}}$) and the risk-neutral correlation ($\rho^{\mathbb{Q}}$), i.e.

$$CRP_{ij,t} = \rho^{\mathbb{P}}_{ij,t} - \rho^{\mathbb{Q}}_{ij,t}. \tag{2}$$

The objective correlation can be determined by the sample time series or by any parametric or non-parametric model. Please note that the use of historical data may be a possible source of inertia for the resulting estimators. However, as we rely on observable market prices to determine the correlation risk premium, we expect a timely readjustment of implied correlation values to changing market conditions.

Assume that correlation risk premia are driven by a single factor ρ_t through

$$CRP_{ij,t} = \rho_t \times \frac{(\sigma^{\mathbb{Q}}_{M,t})^2}{\sigma^{\mathbb{Q}}_{i,t} \times \sigma^{\mathbb{Q}}_{j,t}}. \tag{3}$$

The specification implies that correlations rise with increasing market volatility, ceteris paribus. This behavior can be confirmed empirically: It is well-known that market volatilities rise along with asset correlations during equity market downturns. Thus, when we observe expensive index options, that is, investors seek downside protection, our estimators for the credit risk premia rise and we arrive at lower estimations for our diversification potential.

To restrict the absolute correlation values to be below or equal to 1, we have to standardize our estimations by

$$\rho^{\mathbb{Q}}_{i,t} = \frac{\rho^{\mathbb{Q}}_{ij,t}}{\sqrt{\rho^{\mathbb{Q}}_{ii,t}}\sqrt{\rho^{\mathbb{Q}}_{jj,t}}}. \tag{4}$$

Taken together, the estimator reads

$$\left(\sigma^{\mathbb{Q}}_{M,t}\right)^2 = \sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j \sigma^{\mathbb{Q}}_{i,t}\sigma^{\mathbb{Q}}_{j,t} \frac{\rho^{\mathbb{P}}_{ij,t} - \rho_t \times \frac{(\sigma^{\mathbb{Q}}_{M,t})^2}{\sigma^{\mathbb{Q}}_{i,t}\times\sigma^{\mathbb{Q}}_{j,t}}}{\sqrt{\rho^{\mathbb{P}}_{ii,t} - \rho_t \times \frac{(\sigma^{\mathbb{Q}}_{M,t})^2}{(\sigma^{\mathbb{Q}}_{i,t})^2}}\sqrt{\rho^{\mathbb{P}}_{jj,t} - \rho_t \times \frac{(\sigma^{\mathbb{Q}}_{M,t})^2}{(\sigma^{\mathbb{Q}}_{j,t})^2}}}. \tag{5}$$

The value of $\rho_t \leq 0$ is found numerically.

This specification has many desirable properties.[5] First, it incorporates a negative correlation risk premium, found in Driessen et al. (2009). Furthermore, it is larger in absolute terms for low or negatively correlated asset pairs.[6] Second, the model guarantees that the restriction of Eq. (1) is satisfied under both measures. Third, under mild regularity restrictions, the resulting correlation matrix is symmetric and positive semi-definite.[7]

As a third method of estimating the return correlation matrix, we consider the Shrinkage estimator introduced in Ledoit and Wolf (2003, 2004). It is well-known that the correlation matrix is singular if the estimation window length (T) is smaller than the number of assets (N). Thus, we face the problem of overfitting the correlation matrix with short estimation windows. Extending its length, however, results in higher inertia of the estimator, which means that it is probably not able to adapt to changing market situations quickly enough. Therefore, the idea of the Shrinkage estimator is to impose a factor structure on the correlation estimator by taking a weighted average of the sample correlation matrix with the correlation structure implied by the single-index Market Model. For ease of notation, let $\rho^{\mathbb{P}}$ denote the sample estimator with an estimation window length of 250 calendar days and ρ^{MM} the corresponding correlation matrix implied by the Market Model. The Shrinkage estimator is defined by

$$\rho^{Sh} = (1-\phi)\rho^{\mathbb{P}} + \phi\rho^{MM}, \tag{6}$$

where ϕ denotes the shrinkage intensity parameter. Its optimal value is determined by minimizing the Frobenius norm between the ρ^{Sh} and the true correlation matrix.[8]

2.2 Measuring systematic and unsystematic risk

To investigate idiosyncratic risk premia, we need to disentangle systematic and unsystematic components of the covariance matrix. Studies on the informational content of unsystematic risk usually split total risk ($TR_i(t)$) into a systematic ($SR_i(t)$) and an unsystematic part ($UR_i(t)$), i.e.

$$TR_i(t) = SR_i(t) + UR_i(t), \tag{7}$$

where $TR_i(t)$ is some measure of historical or implied volatility. The systematic component is defined as the predictable volatility part. Unsystematic risk is then calculated by the remaining residual. There are several caveats worth mentioning. First, many asset pricing models used to measure systematic risk do not ensure positivity of the idiosyncratic risk component. These results may be difficult to interpret economically. Second, $SR_i(t)$ is frequently estimated by affine asset pricing models, e.g. by

[5] See Buss and Vilkov (2011).

[6] Empirically, this was found by Krishnan et al. (2009).

[7] Buss and Vilkov (2011) show that the correlation MIDAS matrix estimator is positive semi-definite if and only if the correlation matrix under the physical measure is positive semi-definite and $\rho_t \leq 0$.

[8] For details, see Ledoit and Wolf (2003, 2004).

predictive regressions. Their additive form implies zero correlation between the systematic risk components, an assumption that is typically violated empirically. Third, the results are highly dependent on the identification of the return-driving factors and their correlation, i.e. on the "correct" asset pricing model. However, the number of return-driving factors and their correlation are rarely known ex ante and their impact might change over time.

A promising remedy to these problems is introduced by Laloux et al. (1999). To illustrate the method, let $\mathbf{\Sigma}(t) = \mathbf{S}(t)\rho(t)\mathbf{S}(t)$ denote the covariance matrix estimator, where $\mathbf{S}(t)$ denotes a diagonal matrix, whose elements contain implied volatility values of the single assets, that is, $\sigma_i(t)$. The diagonal elements of $\mathbf{\Sigma}(t)$ serve as our measure for total risk.[9] Our goal is to separate uncorrelated return-driving risk factors, that is, the systematic or predictable part of $\mathbf{\Sigma}(t)$, from random noise, that is, the idiosyncratic part of the correlation matrix. A natural choice is provided by a principal component decomposition of $\mathbf{\Sigma}(t)$:

$$\mathbf{E}'(t)\mathbf{\Sigma}(t)\mathbf{E}(t) \equiv \mathbf{\Lambda}(t), \tag{8}$$

where $\mathbf{\Lambda}(t) \equiv diag(\lambda_1(t), \ldots, \lambda_{N(t)}(t))$ is a diagonal matrix containing the eigenvalues of $\mathbf{\Sigma}(t)$, sorted in decreasing order, and $N(t)$ denotes the number of sample firms in t. The columns of $\mathbf{E}(t) \equiv (\mathbf{e}_1(t), \ldots, \mathbf{e}_{N(t)}(t))'$ are the corresponding eigenvectors. Meucci (2009) points out that they identify a set of $N(t)$ uncorrelated portfolios, the *principal portfolios*, whose returns are decreasingly responsible for the randomness in the market, as the eigenvalues can be interpreted as their portfolio return variances. Borrowing from Random Matrix Theory, Laloux et al. (1999) separate systematic and unsystematic risk by filtering out insignificant eigenvalues, that is, those eigenvalues that are suspected to be driven by random noise.[10] The identification of information-carrying eigenvectors (risk sources) is achieved by analyzing the point estimates for the eigenvalues compared to the eigenvalue spectrum of purely random matrices.[11] Let $T(t)$ denote the estimation window length of the correlation matrix estimator $\rho(t)$ and $N(t)$ be the number of assets, i.e., the dimensionality of the correlation matrix. In the limit, $N(t) \to \infty$, $T(t) \to \infty$, with fixed $Q = \frac{T(t)}{N(t)}$, it can be shown analytically that the distribution $p(\lambda)$ of eigenvalues λ of the random correlation matrix $\rho(t)$ is given by the celebrated Marčenko and Pastur (1967) spectral density

$$p(\lambda) = \frac{Q}{2\pi}\frac{\sqrt{(\lambda_{\max} - \lambda)(\lambda - \lambda_{\min})}}{\lambda}, \tag{9}$$

for $\lambda_{\min} \le \lambda \le \lambda_{\max}$, where

$$\lambda_{\min,\max} = \left(1 + \frac{1}{Q} \pm 2\sqrt{\frac{1}{Q}}\right).$$

A graphical representation for $Q = 2$ is given in Fig. 1.

[9] The outline of this section is based on Meucci (2009).

[10] Examples of studies using Gaussian Random Matrix Theory for financial applications are Laloux et al. (2000), Drożdż et al. (2001), Plerou et al. (2002), and Malevergne and Sornette (2004). The robustness of the eigenvalue analysis is investigated by Rajkovic (2000), as well as by Sharifi et al. (2004).

[11] See also Rajkovic (2000), as well as Sharifi et al. (2004).

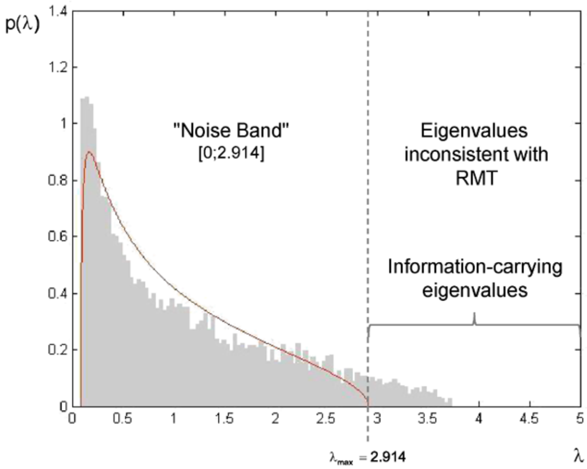

Fig. 1 The distribution of eigenvalues according to random matrix theory. This figure gives a graphical illustration of the Marčenko and Pastur (1967) distribution for $Q = 2$. The largest possible eigenvalue consistent with Random Matrix Theory is $\lambda_{\max} = 2.914$. It separates the "noise band" from the subset of information-carrying eigenvalues

Components with eigenvalues larger than $\lambda_{\max} = 2.914$ cannot be driven by random noise and are thus considered to be information-carrying. Consequently, components with an eigenvalue lower than $\lambda_{\max}$, i.e., those falling into the so-called "noise band", are typically considered to be diversifiable. This simple concept allows us to split $\mathbf{\Sigma}(t)$ into a predictable or information-carrying part representing systematic risk ($\mathbf{\Sigma}^{SR}(t)$) and a part representing idiosyncratic risk ($\mathbf{\Sigma}^{UR}(t)$), which is suspected to be driven by random noise. First, we determine two disjoint eigenvalue subsets by $\lambda \leq \lambda_{\max}$ and $\lambda > \lambda_{\max}$, which are collected in the diagonal matrices

$$\begin{aligned}\mathbf{\Lambda}^{UR}(t) &= diag(\lambda_i : \lambda_i \leq \lambda_{\max}\},\\ \mathbf{\Lambda}^{SR}(t) &= diag(\lambda_i : \lambda_i > \lambda_{\max}\}.\end{aligned} \tag{10}$$

The corresponding eigenvectors are collected in independent eigenvector matrices by

$$\begin{aligned}\mathbf{E}^{UR}(t) &= \left\{\mathbf{e}_i(t)|\forall i : \lambda_i \leq \lambda_{\max}\right\},\\ \mathbf{E}^{SR}(t) &= \left\{\mathbf{e}_i(t)|\forall i : \lambda_i > \lambda_{\max}\right\}.\end{aligned} \tag{11}$$

Systematic and idiosyncratic risk are defined as the diagonal elements of $\mathbf{\Sigma}^{SR}(t) = \mathbf{E}^{SR}(t)\mathbf{\Lambda}^{SR}(t)\mathbf{E}^{SR}(t)'$ and $\mathbf{\Sigma}^{UR}(t) = \mathbf{E}^{UR}(t)\mathbf{\Lambda}^{UR}(t)\mathbf{E}^{UR}(t)'$, respectively.

Our method has several desirable properties. First, it is robust to nonzero correlations between the return-driving factors. Second, the number of factors can vary over time. Third, for $\det \mathbf{\Lambda} \neq 0$, $\mathbf{\Sigma}^{SR}(t)$ and $\mathbf{\Sigma}^{UR}(t)$ are positive definite.

Focardi and Fabozzi (2009) note that eigenvalues and their corresponding return-driving factors have straightforward economic interpretations. Empirically, the portfolio with the largest eigenvalue assigns about equal weights to all portfolio constituents. Therefore, it may be interpreted as the market factor with a similar impact on all sample stocks. Several of the next ranked eigenvalues exhibit positive weights to small stable subsets of the sample space and can be interpreted as specific industry factors.

To investigate the robustness of our results, we implement a method outlined in Meucci (2009) and control for three widely accepted asset pricing models: the classic Market Model (MM), the Fama and French (1993) three-factor model (FF), and the Carhart (1997) model (Mom). Let $\mathbf{A}$ be a $K \times N(t)$ matrix whose rows represent the null space of the factor-mimicking portfolio weights for our three asset pricing models. For $n = K + 1, \dots, N(t)$, we can determine portfolio weights $\mathbf{e}_n$ such that

$$\begin{aligned} \mathbf{e}_n(t) &\equiv \underset{\mathbf{e}(t)'\mathbf{e}(t)\equiv 1}{\arg\max} \left\{\mathbf{e}(t)'\mathbf{\Sigma}(t)\mathbf{e}(t)\right\}, \\ s.t. \quad &\begin{cases} \mathbf{A}\mathbf{e}(t) \equiv \mathbf{0}, \\ \mathbf{e}(t)'\mathbf{\Sigma}(t)\mathbf{e}_j(t) \equiv 0, \quad \forall \mathbf{e}_j(t). \end{cases} \end{aligned} \tag{12}$$

This method is essentially a Principal Component Analysis conditional on the systematic factor components.[12] As before, its resulting eigenvectors and eigenvalues allow splitting $\mathbf{\Sigma}(t)$ into two independent subspaces of systematic ($\mathbf{\Sigma}^{SR}(t)$) and idiosyncratic risk ($\mathbf{\Sigma}^{UR}(t)$).

3 Data description and preparation

3.1 Stock and option data

Our study is based on the S&P 100 index, which contains the largest and most established U.S. stocks with exchange-listed options. Although studies on portfolio selection are typically based on the S&P 500, our choice of the S&P 100 is motivated by several considerations. First, restricting the asset universe to 100 companies at each point in time makes the identification of the correlation matrix more efficient. It is well-known in the literature that the estimator for the correlation matrix is always singular if the number of assets exceeds the number of observations per asset. Second, for each index constituent of the S&P 100, we observe highly liquid option prices. This should reduce biases deriving from illiquid derivative quotes. Third, and probably most importantly, we performed our exercise on the constituents of the DJ Industrial Average, whose constituents consist of the 30 major U.S. companies. This choice would have highly reduced the volatility of our exogenous variables. As such, we consider the S&P 100 to be a good compromise between having an index of high dimensionality with high computational effort and having an index with only low dimensionality that includes very similar companies with respect to the implied return distribution properties.

Our sample stretches from January, 4th, 1996, to October, 30th, 2010. We identify all S&P 100 constituents from the Compustat North American index constituents file. To calculate asset returns, we merge our dataset with CRSP data. These time series account for share splits, dividend payments, and other price factor changes. Then, we merge this dataset with the OptionMetrics volatility surface files, which

[12] For details on the estimation, see Meucci (2009).

contain implied volatility data for options with standardized delta values and time to maturities. Although most variables are available at daily frequency, we decide to base our study on monthly observations of options with a time to maturity of 30 days. This choice is driven by three considerations. First, as options expire on the Saturday following the third Friday of each month, we re-balance our portfolio on Thursdays. This ensures that we do not encounter problems with non-trading days. Second, options with a time to maturity of about 30 calendar days turn out to be highly liquid. Therefore, we expect them to carry a high degree of information. Third, we do not encounter problems deriving from the interpolation scheme over time to obtain a standardized value of 30 days left to maturity.

Let $\sigma_i^{\pm 50}(t)$ denote the implied volatility of asset i's calls and puts with a delta of 50. We determine at-the-money implied volatility of asset i by $\sigma_i^{ATM}(t) = 0.5 \cdot (\sigma_i^{50}(t) + \sigma_i^{-50}(t))$. The corresponding measure of implied skewness is determined by standardized 40-delta risk reversals, i.e., $skew_i(t) = (\sigma_i^{40}(t) - \sigma_i^{-40}(t))/\sigma_i^{ATM}(t)$.[13]

3.2 Construction of the Carhart factor-mimicking portfolios

Our study requires monthly factor returns starting on Thursday. As the available data on Kenneth French's website offer only series that do not match our time periods, we need to calculate our own time series. We use Thursday's closing values as the relevant share prices. All firms contained in the S&P 100 constitute the relevant sample for the construction of the factor premia. To determine the market-to-book ratio (*MTB*), we merge our dataset with the Compustat database. *MTB* is defined as [market value of equity (ME) + total debt (D)]/total assets (A). Total debt is the sum of total liabilities, preferred stock, and convertible debt minus deferred taxes. The market value of equity is calculated as the number of outstanding shares multiplied by the closing stock price.

There is a broad consensus on how to determine the market factor. Following Fama and French (1993), and Carhart (1997), it is calculated as the excess return of a market-weighted portfolio over the risk-free rate (defined as the monthly Treasury Bill yield). The premia of the remaining three factors are determined by zero-investment portfolios as follows. Each stock in the database is ranked by its monthly median market capitalization (defined as closing price multiplied by the number of outstanding shares) and sorted into one of the two groups "Big" (B) and "Small" (S). At the same time, we assign each stock to two groups by its market-to-book ratio ("High" (H) and "Low" (L)) and by its past one-year return ("Up" (U) and "Down" (D)). Then, we calculate value-weighted returns for each of the eight sub-portfolios. Following Fama and French (1993), we determine the desired premia by the returns of the zero-investment factor-mimicking portfolios SMB ("Small minus Big"), HML ("High minus Low"), and UMD ("Up minus Down").

[13] This definition of implied skewness might seems a bit crude at first sight. However, more sophisticated measures, such as Bakshi et al. (2003), typically require liquid option quotes for far out-of-the-money strike prices. These are usually unavailable for equity options.

4 Empirical results

4.1 Estimation of the return covariance matrix and factor return volatilities

This section highlights our empirical results. In the first step, we investigate the informational content of our correlation and return volatility predictors. *Panel A* of Table 1 gives root mean squared errors for our three correlation predictors: The historical return correlation with a 250-day estimation window ($\rho^{\mathbb{P}}$), the corresponding HETIC correlation of Buss and Vilkov (2011) (ρ^{HET}), and the Shrinkage estimator introduced in Ledoit and Wolf (2003, 2004) (ρ^{Sh}).

Panel B illustrates the corresponding results for predictive regressions of the form

$$\rho_{ij}(t, t+30) = \alpha + \gamma \rho_{ij}(t) + \tilde{\varepsilon}_{ij}(t),$$

where $\rho_{ij}(t, t+30)$ denotes the realized equity return correlation of assets i and j between t and $t+30$, measured by daily closing prices. $\rho_{ij}(t)$ is its time-t predictor and $\tilde{\varepsilon}_{ij}(t)$ is the error term. The standard errors are corrected by the Newey and West (1987) estimator for panel data with a lag length of up to two periods.

Regarding the complete sample period, the predictive power of the three estimators is of comparable magnitude with root mean squared error values ranging from 0.2469 for the HETIC estimator to 0.2507 for historical correlation values. The coefficients of our predictive regressions are all highly significant. We observe adjusted R^2 values of 0.1812 for $\rho^{\mathbb{P}}$ to 0.1935 for ρ^{HET}. Our results are in line with the findings of DeMiguel et al. (2009), who investigate the diversification potential for different volatility and correlation estimators. Interestingly, they find that correlation estimations are of only minor importance for diversification benefits. Our results show that this may be caused by a similar forecasting accuracy.[14]

Similar conclusions can be drawn from investigating the corresponding values of the subperiods from January 1996 to March 2008 and from April 2008 to October 2010. As expected, we find a higher forecasting accuracy for the period before the subprime crisis. We observe very similar root mean squared errors of 0.2445 for ρ^{HET} to 0.2485 for $\rho^{\mathbb{P}}$. The corresponding R^2 values range from 0.1366 for the historical estimator to 0.1547 for the HETIC predictor of Buss and Vilkov (2011). For the subperiod of the subprime crisis of April 2008 to October 2010, we observe that the root mean squared error of ρ^{HET} has a relatively low value of 0.2512, compared to 0.2604 and 0.2617 for $\rho^{\mathbb{P}}$ and ρ^{Sh}, respectively. This finding shows that the HETIC estimator is able to incorporate valuable information from the option market for correlation prediction purposes.

Panels C and *D* of Table 1 give results on the root mean squared errors and predictive regressions for equity return volatilities. We include four different predictors: historical return volatilities with estimation windows of 30 ($\sigma_{30}^{\mathbb{P}}$), 250 ($\sigma_{250}^{\mathbb{P}}$), and 750 ($\sigma_{750}^{\mathbb{P}}$) calendar days, measured by the returns of daily closing prices. In addition, we consider option-implied volatilities with a corresponding time to expiry of 30 days (σ^{ATM}). In general, root mean squared errors rise with the length of the estimation

[14] See also Pojarliev and Polasek (2003).

Table 1 Prediction of realized volatility and correlation

		Complete sample	01/1996–03/2008	04/2008–10/2010
Panel A: Root mean squared errors for correlation predictions				
$\rho^{\mathbb{P}}$		0.2507	0.2485	0.2604
ρ^{HET}		0.2469	0.2445	0.2512
ρ^{Sh}		0.2490	0.2462	0.2617
Panel B: Predictive correlation regressions				
$\rho^{\mathbb{P}}$	α	0.1159***	0.1131***	0.1291**
	γ	0.5892***	0.5755***	0.6541***
	R^2	0.1812	0.1366	0.0704
ρ^{HET}	α	0.0410***	0.0358***	0.0658
	γ	0.7651***	0.7538***	0.8188***
	R^2	0.1935	0.1547	0.0923
ρ^{Sh}	α	0.0829***	0.0805***	0.0943
	γ	0.7181***	0.7193***	0.7127***
	R^2	0.1911	0.1481	0.0682
Panel C: Root mean squared errors for volatility predictions				
$\sigma_{30}^{\mathbb{P}}$		0.1853	0.1631	0.2647
$\sigma_{250}^{\mathbb{P}}$		0.1913	0.1649	0.3089
$\sigma_{750}^{\mathbb{P}}$		0.2206	0.1714	0.3730
σ^{ATM}		0.1595	0.1356	0.2410
Panel D: Predictive volatility regressions				
$\sigma_{30}^{\mathbb{P}}$	α	0.0838***	0.0960***	0.0774***
	γ	0.7474***	0.6975***	0.8059***
	R^2	0.5351	0.4637	0.6223
$\sigma_{250}^{\mathbb{P}}$	α	0.0597***	0.0675***	0.0232
	γ	0.7788***	0.7484***	0.9216***
	R^2	0.3763	0.4022	0.3128
$\sigma_{750}^{\mathbb{P}}$	α	0.0473***	0.0608***	−0.0164
	γ	0.8072***	0.7361***	1.1421***
	R^2	0.1982	0.2871	0.0604
σ^{ATM}	α	0.0099	0.0207**	−0.0410**
	γ	1.1009***	1.0606***	1.2893***
	R^2	0.5917	0.5535	0.6249

This table gives results on the predictive power of different correlation measures. *Panel A* shows root mean squared errors of the monthly correlation predictions for the complete sample period, as well as for the subperiods 01/1996–03/2008 and 04/2008–10/2010. Our correlation predictors are the historical daily correlation ($\rho^{\mathbb{P}}$) based on the closing prices with a window length of 250 calendar days. In addition, we also consider the corresponding HETIC correlation estimator (ρ^{HET}), of Buss and Vilkov (2011) using a historical time series with a window length of 250 days, as well as the shrinkage estimator (ρ^{Sh}) of Ledoit and Wolf (2003, 2004). We shrink the sample estimate with a window length of 250 days toward the corresponding correlation matrix implied by the Market Model. The optimal shrinkage density is found by minimizing the Frobenius norm between the shrinkage estimator and the true correlation matrix. Our sample extends from January 4th, 1996, to October 30th, 2010. Our option dataset is obtained from OptionMetrics. The corresponding values for share prices are from the Compustat database.
The results for our predictive regressions are given in *Panel B*. They are of the form

$$\rho_{ij}(t, t+30) = \alpha + \gamma \rho_{ij}(t) + \tilde{\varepsilon}_{ij}(t),$$

where $\rho_{ij}(t, t+30)$ is the realized equity return correlation of assets i and j between t and $t+30$, measured by daily closing prices. Its time-t predictor is denoted by $\rho_{ij}(t)$. Corresponding values for our return volatility predictions are given in *Panels C* and *D*, respectively. We consider historical return volatility estimators with estimation window lengths of 30 ($\sigma_{30}^{\mathbb{P}}$), 250 ($\sigma_{250}^{\mathbb{P}}$), and 750 ($\sigma_{750}^{\mathbb{P}}$) calendar days, as well as at-the-money implied volatility (σ^{ATM}) derived from options with a delta of 50 and a time to maturity of 30 calendar days. Standard errors are corrected for serial correlation up to two lags by the Newey and West (1987) estimator

***, **, and * denote significance at the 99, 95, and 90 percent level, respectively window. Values range from 0.1853 for $\sigma_{30}^{\mathbb{P}}$ to 0.2206 for $\sigma_{750}^{\mathbb{P}}$. A similar pattern can be seen in the corresponding regression results. Our finding is clearly driven by inertia of the volatility predictors. This is especially apparent during the April 2008 to October 2010 subperiod. The volatility shock in the course of the Lehman collapse is most quickly incorporated by $\sigma_{30}^{\mathbb{P}}$. It has a root mean squared error of 0.2647, which is low compared to 0.3730 for $\sigma_{750}^{\mathbb{P}}$. Regarding implied volatility, we find that the predictive power is clearly higher than for estimators based on historical volatility. Again, the estimator is better able to capture changing market situations very quickly. This result is in line with previous findings in the literature. It is robust to both subperiods.

Next, we investigate the predictability of the return volatilities for the factor-mimicking portfolios. As before, let $\mathbf{S}(t)$ be a matrix of zeros, whose diagonal elements include the implied volatilities $\sigma_i^{ATM}(t)$ of sample options with a remaining time to maturity of 30 calendar days. In addition, let $\mathbf{w}_F(t)$ be the vector of asset weights for the zero-investment factor-mimicking portfolio, where $F \in \{SMB, HML, UMD\}$. The estimators for the factor volatility can be written as $\sigma_F(t) = \sqrt{\mathbf{w}_F'(t)\mathbf{S}(t)\rho(t)\mathbf{S}(t)\mathbf{w}_F(t)}$, where $\rho(t)$ denotes the estimator of the correlation matrix. Again, we consider the historical correlation with an estimation window length of 250 calendar days, the HETIC correlation estimator of Buss and Vilkov (2011), and the shrinkage estimator introduced in Ledoit and Wolf (2003, 2004). Root mean squared errors and results of the predictive regressions are given in *Panels A* and *B* of Table 2.

In general, we observe that all three estimators achieve the highest predictive power for the SMB return volatility. Root mean squared errors range from 0.0353 for ρ^{HET} to 0.0425 for ρ^{Sh}. Their values strongly increase for the HML factor volatility, ranging from 0.0575 (ρ^{HET}) to 0.0756 (ρ^{Sh}). These findings are robust to both subsamples, especially for the period of the financial crisis. The highest root mean squared errors can be found for the prediction of the UMD factor volatility. We observe values ranging between 0.0646 ($\rho^{\mathbb{P}}$) to 0.0801 (ρ^{Sh}).

Next, we investigate the results of our Fama and MacBeth (1973) regressions of systematic and idiosyncratic risk, which are given in Table 3. They are of the form

$$r_i(t, t+30) = \alpha + \gamma_{UR} UR_i(t) + \gamma_{SR} SR_i(t) + \tilde{\varepsilon}_i(t). \tag{13}$$

In general, we find positive and highly significant impacts of both systematic and unsystematic risk on future stock returns. Interestingly, adjusted determination coefficients are of comparable magnitude for the unconstrained model, as well as the three constrained estimators. Furthermore, all regressions have very similar parameter estimates. This finding shows that the unconstrained model typically defines the risk of common asset pricing models as systematic. During the January 1996 to March 2008 subperiod, the link between idiosyncratic risk and future equity returns is weak and in some cases even insignificant. However, we can find a strong impact of systematic risk for all models. Different conclusions can be drawn for the April 2008 to October 2010 subperiod, in which we find that unsystematic risk has a very strong and highly significant impact on future returns.

It is well-known in the literature that Fama and MacBeth (1973) analyses of risk measures face the "errors in variables" problem. This means that our results may

Table 2 Prediction of realized volatility for factor portfolios

		SMB			HML			UMD		
		Complete sample	01/1996–03/2008	04/2008–10/2010	Complete sample	01/1996–03/2008	04/2008–10/2010	Complete sample	01/1996–03/2008	04/2008–10/2010
Panel A: Root mean squared errors for factor return volatilities										
$\rho^{\mathbb{P}}$		0.0365	0.0350	0.0430	0.0576	0.0458	0.0954	0.0646	0.0556	0.0964
ρ^{HET}		0.0353	0.0354	0.0416	0.0575	0.0456	0.0934	0.0657	0.0580	0.0940
ρ^{Sh}		0.0425	0.0421	0.0442	0.0756	0.0597	0.1262	0.0801	0.0688	0.1200
Panel B: Predictive regressions for factor return volatilities										
$\rho^{\mathbb{P}}$	α	−0.0113	−0.0099	−0.0129	−0.0222***	−0.0020	−0.0436	−0.0140	−0.0043	−0.0198
	γ	1.1051***	1.0872***	1.1434***	1.1901***	1.0279***	1.3382***	1.1942***	1.1177***	1.2677***
	R^2	0.6772	0.6545	0.7314	0.7840	0.6955	0.8234	0.7633	0.7015	0.8110
ρ^{HET}	α	−0.0074	−0.0109	−0.0047	−0.0124	0.0020	−0.0286	−0.0073	−0.0042	−0.0084
	γ	1.1197***	1.1619***	1.0581***	1.1834***	1.0636***	1.2853***	1.1951***	1.1699***	1.2145***
	R^2	0.6871	0.6485	0.7374	0.7955	0.6966	0.8265	0.7614	0.6843	0.8129
ρ^{Sh}	α	−0.0014	−0.0023	−0.0042	−0.0109	0.0087	−0.0147	−0.0011	0.0104	0.0001
	γ	1.2164***	1.2377***	1.1965***	1.4063***	1.1921***	1.5427***	1.2990***	1.1816***	1.4014***
	R^2	0.6713	0.6328	0.7279	0.7623	0.6302	0.8141	0.7234	0.6415	0.8054

This table shows results on the predictive power for the factor return volatilities of the Fama and French (1993) model, and the Carhart (1997) model. Every sample month, we determine the zero-investment factor-mimicking portfolios SMB ("Small minus Big"), HML ("High minus Low"), and UMD ("Up minus Down"). Let $\mathbf{w}_F(t)$ be the vector of the factor portfolio weights for $F \in \{SMB, HML, UMD\}$ and $\mathbf{S}(t)$ denote a matrix of zeros, whose diagonal contains the implied volatilities $\sigma_i^{ATM}(t)$ of sample options with a remaining time to maturity of 30 calendar days. The estimators for our factor return volatilities can be written as

$$\sigma_F(t) = \sqrt{\mathbf{w}'_F(t)\mathbf{S}(t)\rho(t)\mathbf{S}(t)\mathbf{w}(t)},$$

where $\rho(t)$ denotes the estimator of the correlation matrix. Our study includes three different predictors for the correlation matrix. We use the historical daily correlation ($\rho^{\mathbb{P}}$), based on the closing prices with a window length of 250 calendar days. In addition, we also consider the corresponding HETIC correlation estimator (ρ^{HET}), of Buss and Vilkov (2011) using a historical time series with a window length of 250 days, as well as shrinkage estimator (ρ^{Sh}) of Ledoit and Wolf (2003, 2004). We shrink the sample estimate with a window length of 250 days toward the corresponding correlation matrix implied by the Market Model. The optimal shrinkage density is found by minimizing the Frobenius norm between the shrinkage estimator and the true correlation matrix. *Panel A* gives root mean squared errors for the complete sample period from 01/1996 to 10/2010, as well as the two subperiods 01/1996 to 03/2008 and 04/2008 to 10/2010. Results of our predictive regressions are given in *Panel B*. These are of the form

$$\sigma_F(t, t+30) = \alpha + \gamma\sigma_F(t) + \tilde{\varepsilon}_F(t),$$

where $\sigma_F(t, t+30)$ is the realized factor return volatility between t and $t+30$, measured by daily closing prices. The standard errors are corrected for serial correlation up to two lags by the Newey and West (1987) estimator

***, **, and * denote significance at the 99, 95, and 90 percent level, respectively

Table 3 The effects of systematic and unsystematic implied volatility—Fama–MacBeth regression results

	Unconstrained model			MM			FF			Mom		
	$\rho^{\mathbb{P}}$	ρ^{HET}	ρ^{Sh}	$\rho^{\mathbb{P}}$	ρ^{HET}	ρ^{Sh}	$\rho^{\mathbb{P}}$	ρ^{HET}	ρ^{Sh}	$\rho^{\mathbb{P}}$	ρ^{HET}	ρ^{Sh}
Panel A: Complete sample												
α	−0.0166**	−0.0162**	−0.0162**	−0.0166**	−0.0162**	−0.0163**	−0.0163**	−0.0165**	−0.0165**	−0.0163**	−0.0165**	−0.0167**
γ_{UR}	0.0358**	0.0418**	0.0502***	0.0367**	0.0416*	0.0499***	0.0425*	0.0424**	0.0504***	0.0395**	0.0390**	0.0507***
γ_{SR}	0.0652***	0.0555**	0.0526*	0.0640**	0.0555**	0.0536*	0.0582**	0.0578**	0.0540*	0.0600***	0.0615**	0.0549*
R^2	0.0064	0.0061	0.0061	0.0063	0.0060	0.0060	0.0067	0.0062	0.0060	0.0066	0.0062	0.0060
Panel B: 01/1996–03/2008												
α	−0.0127	−0.0128	−0.0126	−0.0127	−0.0128	−0.127	−0.0122	−0.0129	−0.0130	−0.0123	−0.0130	−0.0133
γ_{UR}	0.0138	0.0352*	0.0345**	0.0149	0.0350*	0.0342**	0.0148	0.0240	0.0357***	0.0183	0.0248	0.0372**
γ_{SR}	0.0755***	0.0564**	0.0628**	0.0741***	0.0564**	0.0639**	0.0717***	0.0678**	0.0629**	0.0691***	0.0681**	0.0628**
R^2	0.0035	0.0039	0.0032	0.0033	0.0037	0.0033	0.0033	0.0034	0.0033	0.0034	0.0034	0.0033
Panel C: 04/2008–10/2010												
α	−0.0414*	−0.0377*	−0.0391*	−0.0414*	−0.0377*	−0.0391*	−0.0419*	−0.0392*	−0.0392*	−0.0412*	−0.0394*	−0.0392*
γ_{UR}	0.1218**	0.0446**	0.0955**	0.1217**	0.0449**	0.0954**	0.1506**	0.0972***	0.0977**	0.1255**	0.0929***	0.0946**
γ_{SR}	0.0339*	0.0785***	0.0383*	0.0339*	0.0782***	0.0384*	0.0247**	0.0388***	0.0377*	0.0318*	0.0448***	0.0408*
R^2	0.0196	0.0184	0.0168	0.0194	0.0182	0.0166	0.0223	0.0172	0.0166	0.0200	0.0174	0.0166

This table illustrates the results of our regression analysis on systematic (SR) and unsystematic (UR) risk. Total risk is defined by the diagonal elements of $\mathbf{\Sigma}(t) = \mathbf{S}(t)\rho(t)\mathbf{S}(t)$, where $\mathbf{S}(t)$ denotes a matrix of zeros, whose diagonal elements contains the implied volatilities $\sigma_i(t)$ of sample firm options with a remaining time to maturity of 30 calendar days. The correlation structure $\rho(t)$ between the single assets is estimated by the historical correlation based on daily closing prices with an estimation window length of 250 days ($\rho^{\mathbb{P}}$). In addition, we also consider the corresponding HETIC correlation estimator (ρ^{HET}), of Buss and Vilkov (2011) using a historical time series with a window length of 250 days as well as shrinkage estimator (ρ^{Sh}) of Ledoit and Wolf (2003, 2004). We shrink the sample estimate with a window length of 250 days toward the corresponding correlation matrix implied by the single-factor Market Model. The optimal shrinkage density is found by minimizing the Frobenius norm between the shrinkage estimator and the true correlation matrix. Conducting a Principal Component Analysis (PCA) on $\mathbf{\Sigma}(t)$ allows to split total risk into two independent subspaces. The first ($\mathbf{\Sigma}^{SR}(t)$) is spanned by those eigenvectors, whose corresponding eigenvalues are above the set of theoretical eigenvalues attainable by Gaussian random matrices. The second subspace ($\mathbf{\Sigma}^{UR}(t)$) is spanned by the remaining eigenvectors. The implied risk components of each firm i ($SR_i(t)$ and $UR_i(t)$) are defined by the ith diagonal elements of the corresponding matrices. We estimate the equation

$$r_i(t, t+1) = \alpha + \gamma_{UR} UR_i(t) + \gamma_{SR} SR_i(t) + \tilde{\varepsilon}_i(t),$$

where $\tilde{\varepsilon}_i(t)$ denotes the error term. *Panel A* shows results for the complete sample period from 01/1996, until 10/2010. Corresponding values for the subperiods from 01/1996 through 03/2003, and from 04/2008 through 10/2010, can be found in *Panel B* and *C*, respectively. Corresponding values for our Fama–MacBeth regressions controlling for implied skewness are given in *Panels D*, *E*, and *F*. Implied skewness is calculated as the 40 delta risk reversal of options with a remaining time to expiry of 30 calendar days. Our standard errors are determined by the Newey and West (1987) estimator for panel data with a pre-specified lag length of two periods

***, **, and * indicate significance at the 99, 95, and 90 percent level, respectively be caused by noisy variable estimations. The usual remedy for this problem in the literature is to form equity portfolios in the hope that the potential biases diversify in the portfolio context. We follow this route next. Each month, we sort all sample firms based on $UR_i(t)$ and assign them to one of five equally weighted quintile portfolios.[15] Then, we calculate the return time series of the portfolio with the highest ("Hi") and the lowest ("Lo") unsystematic risk, as well as the time series of the corresponding difference return ("Hi–Lo"). Results are given in *Panel A* of Table 4.

In general, we find positive premia for unsystematic risk in all three asset pricing models. The average monthly premium mounts to 0.14 percent. This value is statistically and economically significant. Interestingly, we do not find significant difference premia for the April 2008 to October 2010 subperiod. The corresponding analysis for systematic risk is illustrated in *Panel B* of Table 4. Here, we see a slightly different picture, compared to the results of our Fama and MacBeth (1973) regressions. Again, we find positive difference premia. These are significant for the first subperiod; however, during the subperiod of the financial crisis, our finding is not significant. The results of our sorting exercise on implied skewness are outlined in *Panel C*. We find positive return premia on implied skewness for the complete sample, with an average annual premium of 0.62 percent. Similar results are found for the January 1996 to March 2008 subperiod. However, we cannot find comparable premia for the April 2008 to October 2010 subperiod.

In the literature, the usual way of explaining unsystematic premia is to extend the pricing kernel, for example, by higher return distribution moments. This leads to asset pricing models in which higher moments of the implied return distribution have an impact on asset returns (Vanden 2006).[16] To investigate the relationship between premia on unsystematic risk and implied equity return skewness, we conduct a double-sorting analysis. Each month, we sort the sample firms with respect to UR_i and assign them to tercile portfolios. For each portfolio, we define three equally weighted subportfolios based on a sorting on implied return skewness. Average returns, as well as average difference returns are given in Table 5.

To some extent, the results indicate that the premia for bearing unsystematic risk can be explained by implied skewness. We find that unsystematic risk has positive premia only for those assets that have positive implied skewness. Interestingly, we cannot find similar results for the April 2008 to October 2010 subperiod.

5 Robustness check—the distribution of eigenvalues in finite samples

In the following, we are concerned with the fact that the validity of the Marčenko and Pastur (1967) spectral density in Eq. (9) may be sensitive to the assumption of $T \to \infty$ and $N \to \infty$.[17] As our correlation matrix estimators $\rho(t)$ are based on finite

[15] Our results are robust to forming value-weighted portfolios.

[16] Among the most influential models are Kraus and Litzenberger (1976) and Harvey and Siddique (2000). However, there exists a vast literature on (co-)skewness models.

[17] For the distributional properties of $\lambda_{\max}$, see Geman (1980), Baik et al. (2005), and the references therein.

Table 4 The individual effects of systematic and unsystematic implied volatility

		Complete sample			01/1996–03/2008			04/2008–10/2010		
		Lo	Hi	Hi–Lo	Lo	Hi	Hi–Lo	Lo	Hi	Hi–Lo
Panel A: Unsystematic risk										
unconst	$\rho^{\mathbb{P}}$	0.0881***	0.2161***	0.1281*	0.1035***	0.1935***	0.0901*	0.0151	0.3234	0.3082
	ρ^{HET}	0.0761***	0.2202***	0.1441**	0.0906***	0.1983***	0.1077*	0.0074	0.3240	0.3167
	ρ^{Sh}	0.0693**	0.2149***	0.1456**	0.0781**	0.1963***	0.1182**	0.0275	0.3031	0.2756
MM	$\rho^{\mathbb{P}}$	0.0872***	0.2147***	0.1275*	0.1021***	0.1920***	0.0899*	0.0164	0.3224	0.3060
	ρ^{HET}	0.0759***	0.2212***	0.1452**	0.0913***	0.1995***	0.1082*	0.0030	0.3240	0.3211
	ρ^{Sh}	0.0702**	0.2149***	0.1447**	0.0792**	0.1963***	0.1171**	0.0275	0.3031	0.2756
FF	$\rho^{\mathbb{P}}$	0.0886***	0.2152***	0.1266**	0.1063***	0.1952***	0.0889	0.0047	0.3102	0.3056
	ρ^{HET}	0.0753***	0.2085***	0.1332*	0.0921***	0.1857***	0.0936*	−0.0043	0.3165	0.3208
	ρ^{Sh}	0.0711**	0.2115***	0.1404**	0.0807**	0.1945***	0.1137*	0.0253	0.2923	0.2670
Mom	$\rho^{\mathbb{P}}$	0.0857***	0.2187***	0.1330**	0.1018***	0.1987***	0.0969*	0.0092	0.3135	0.3043
	ρ^{HET}	0.0744***	0.2173***	0.1429**	0.0918***	0.1908***	0.0989*	−0.0082	0.3433	0.3515
	ρ^{Sh}	0.0751**	0.2087***	0.1335**	0.0865***	0.1906***	0.1041*	0.0215	0.2945	0.2731
Panel B: Systematic risk										
unconst	$\rho^{\mathbb{P}}$	0.0696**	0.2292***	0.1596**	0.0737**	0.2199***	0.1463**	0.0503	0.2732	0.2229
	ρ^{HET}	0.0737***	0.2324***	0.1587**	0.0791***	0.2227***	0.1436**	0.0484	0.2783	0.2299
	ρ^{Sh}	0.0912***	0.2205***	0.1293*	0.1032**	0.2056***	0.1023*	0.0056	0.0067	0.0011
MM	$\rho^{\mathbb{P}}$	0.0698**	0.2278***	0.1581**	0.0739**	0.2187***	0.1448**	0.0501	0.2711	0.2210
	ρ^{HET}	0.0752***	0.2289***	0.1537**	0.0808***	0.2190***	0.1382**	0.0483	0.2758	0.2276
	ρ^{Sh}	0.0909***	0.2217***	0.1308*	0.1029***	0.2070***	0.1041*	0.0341	0.2914	0.2573
FF	$\rho^{\mathbb{P}}$	0.0744***	0.2324***	0.1580**	0.0805***	0.2207***	0.1402**	0.0458	0.2883	0.2426
	ρ^{HET}	0.0746***	0.2280***	0.1534**	0.0802***	0.2154***	0.1351**	0.0478	0.2877	0.2398
	ρ^{Sh}	0.0944***	0.2172***	0.1228*	0.1058***	0.2032***	0.0974	0.0405	0.2833	0.2428
Mom	$\rho^{\mathbb{P}}$	0.0767***	0.2330***	0.1563**	0.0835***	0.2210***	0.1375**	0.0446	0.2902	0.2456
	ρ^{HET}	0.0736***	0.2160***	0.1424**	0.0798***	0.2032***	0.1234**	0.0444	0.2767	0.2322
	ρ^{Sh}	0.0923***	0.2135***	0.1212*	0.1031***	0.1988***	0.0957	0.0410	0.2831	0.2421
Panel C: Implied skewness										
		0.0952**	0.1465***	0.0513**	0.0957**	0.1645***	0.0688***	0.0927	0.0611	−0.0316

This table shows average monthly returns of portfolios sorted into quintiles on the basis of systematic (SR) and unsystematic (UR) implied variance. Total risk is defined by the diagonal elements of $\mathbf{\Sigma}(t) = \mathbf{S}(t)\rho(t)\mathbf{S}(t)$, where $\mathbf{S}(t)$ denotes a matrix of zeros, whose diagonal elements contains the implied volatilities $\sigma_i(t)$ of sample firm options with a remaining time to maturity of 30 calendar days. The correlation structure $\rho(t)$ between the single assets is estimated by the historical correlation based on daily closing prices with an estimation window length of 250 days ($\rho^{\mathbb{P}}$). In addition, we also consider the corresponding HETIC correlation estimator (ρ^{HET}), of Buss and Vilkov (2011) using a historical time series with a window length of 250 days as well as shrinkage estimator (ρ^{Sh}) of Ledoit and Wolf (2003, 2004). We shrink the sample estimate with a window length of 250 days toward the corresponding correlation matrix implied by the single-factor Market Model. The optimal shrinkage density is found by minimizing the Frobenius norm between the shrinkage estimator and the true correlation matrix. Conducting a Principal Component Analysis (PCA) on $\mathbf{\Sigma}(t)$ allows to split total risk into two independent subspaces. The first ($\mathbf{\Sigma}^{SR}(t)$) is spanned by those eigenvectors, whose corresponding eigenvalues are above the set of theoretical eigenvalues attainable by Gaussian random matrices. The second subspace ($\mathbf{\Sigma}^{UR}(t)$) is spanned by the remaining eigenvectors. The implied risk components of each firm i ($SR_i(t)$ and $UR_i(t)$) are defined by the ith diagonal elements of the corresponding matrices.

The table illustrates average returns for the quintile portfolios, formed on the basis of independent sortings on unsystematic (*Panel A*) and systematic firm risk (*Panel B*). Our sample period extends from 01/1996 to 10/2010. In addition, the table provides results for the subperiods from 01/1996 through 03/2008, and from 04/2008 through 10/2010. Significance on the average returns of the difference portfolios (Hi–Lo) is determined by the Newey and West (1987) estimator with a pre-specified lag length of two periods

***, **, and * indicate significance at the 99, 95, and 90 percent level, respectively

Table 5 The effect of implied skewness and unsystematic risk

Skewness	UR				UR				UR			
	$\rho^{\mathbb{P}}$				ρ^{HET}				ρ^{Sh}			
	Lo	Mid	Hi	Hi–Lo	Lo	Mid	Hi	Hi–Lo	Lo	Mid	Hi	Hi–Lo
Panel A: Complete sample												
Lo	0.0779**	0.1116***	0.0880***	0.0102	0.0681**	0.1191***	0.0765**	0.0085	0.0661***	0.1199***	0.0985***	0.0324
Mid	0.0952**	0.1553***	0.1708***	0.0756***	0.1000**	0.1513***	0.1891***	0.0890***	0.1094***	0.1445***	0.1597***	0.0503*
Hi	0.1943***	0.2314***	0.2370***	0.0426*	0.2004***	0.2268***	0.2322***	0.0319**	0.1933***	0.2323***	0.2372***	0.0493*
Hi–Lo	0.1165**	0.1198*	0.1490***		0.1323**	0.1077*	0.1557***		0.1272**	0.1123*	0.1387***	
Panel B: 01/1996–03/2008												
Lo	0.0895***	0.1306***	0.1050***	0.0155	0.0822***	0.1361***	0.0959***	0.0137	0.0710**	0.1353***	0.1116***	0.0405
Mid	0.1100***	0.1618***	0.1885***	0.0785***	0.1091***	0.1618***	0.2020***	0.0928***	0.1240***	0.1500***	0.1660***	0.0421
Hi	0.1707***	0.1636***	0.2367***	0.0660*	0.1795***	0.1576***	0.2333***	0.0538	0.1760***	0.1701***	0.2496***	0.0736*
Hi–Lo	0.0812*	0.0331	0.1317***		0.0973*	0.0215	0.1374**		0.1050*	0.0348	0.1381*	
Panel C: 04/2008–10/2010												
Lo	0.0226	0.0218	0.0075	−0.0151	0.0013	0.0386	−0.0152	−0.0165	0.0430	0.0470	0.0368	−0.0062
Mid	0.0248	0.1241	0.0870	0.0622	0.0569	0.1014	0.1280	0.0711	0.0401	0.1183	0.1296	0.0896
Hi	0.3066	0.5527*	0.2382	−0.0683	0.2995	0.5548*	0.2275	−0.0720	0.2756	0.5273	0.1782	−0.0974
Hi–Lo	0.2840	0.5308*	0.2307*		0.2982	0.5548*	0.2427*		0.2326	0.4802*	0.1415	

This table shows the results of our double-sorting procedure on implied skewness and unsystematic risk. Each month, we sort our sample firms with respect to unsystematic risk $UR_i(t)$ and assign them to one of three portfolios ("Lo", "Mid", and "Hi"). For each of these portfolios we form three equally weighted subportfolios based on sortings on implied skewness. Our skewness measure is defined as the 40 delta risk reversal of options with a remaining time to expiry of 30 calendar days. *Panel A* shows average monthly returns for each portfolio, as well as the respective difference returns ("Hi–Lo") for the complete sample period from 01/1996 to 10/2010. Corresponding values for the subperiods from 01/1996 through 03/2008 and from 04/2008 through 10/2010 can be found in *Panels B* and *C*, respectively

***, **, and * indicate significance at the 99, 95, and 90 percent level

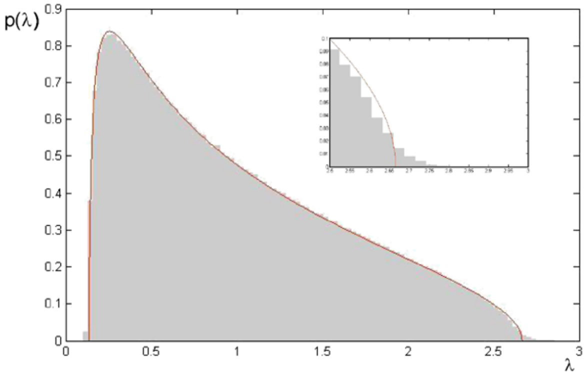

Fig. 2 The distribution of eigenvalues for finite samples. This figure shows a histogram of 10.000 bootstrapped finite-sample eigenvalues, which are determined by uncorrelated random return matrices with $N(t) = 100$ and $T(t) = 200$. The corresponding Marčenko and Pastur (1967) distribution for $Q = 2$ is given in *red*

estimation window lengths ($T < \infty$) and finite sample sizes ($N(t) \leq 100$), we face the problem that the noise band is potentially misspecified. A graphical illustration is given in Fig. 2.

To study finite-sample properties of the eigenvalue distribution, we simulate 10,000 uncorrelated return time series with $N(t) = 100$ and $T(t) = 200$ such that $Q = 2$. Figure 2 is a histogram of the resulting correlation matrix eigenvalues. It shows that the value $\lambda_{max} = 2.914$ of the Marčenko and Pastur (1967) distribution is frequently overshot. This phenomenon is essentially a finite-sample property.[18] However, deriving a functional form for the density of eigenvalues in finite samples is far from trivial.[19] We cannot use standard Central Limit Theorem arguments to apply Gaussian Random Matrix Theory. Therefore, we generate a finite-sample bootstrap distribution of eigenvalues with the following steps. Let the estimator $\rho(t)$ be based on returns that are collected in the $T(t) \times N(t)$ matrix $\mathbf{R}(t) = \{\mathbf{r}_1, \ldots, \mathbf{r}_{N(t)}\}$, where the univariate return time series for sample constituent i is denoted by $\mathbf{r}_i(t) = (r_{i,t-T(t)}(t), \ldots, r_{i,t-1}(t))'$. First, we estimate univariate return distributions $F_i(r, t)$ by

$$F_i(r, t) = \frac{1}{T(t)} \sum_{j=1}^{T(t)} \mathbf{1}_{\{r_{i,t-j}(t) \leq r\}}, \tag{14}$$

where $\mathbf{1}$ denotes the indicator function.

For each asset, we form $T(t) \times 1$ vectors of independent uniformly distributed $U(0, 1)$ random variables $\mathbf{u}_i = (u_{i,1}, \ldots, u_{i,T(t)})$ to construct the pseudo return sample $\hat{\mathbf{R}}(t) = (\hat{\mathbf{r}}_1(t), \ldots, \hat{\mathbf{r}}_{N(t)})$, where

$$\hat{\mathbf{r}}_i(t) = \left(F_i^{-1}(u_{i,1}), \ldots, F_i^{-1}(u_{i,T(t)})\right)'.$$

$\hat{\mathbf{R}}(t)$ can be used to estimate a pseudo correlation matrix $\hat{\rho}(t)$. Then, we determine the eigenvalues of the resulting correlation matrices by standard Principal Component

[18] See Burda et al. (2011).

[19] See Kollo and Ruul (2003).

Analysis. Generating 20,000 pseudo return time series $\hat{\mathbf{R}}(t)$, we have a bootstrapped distribution of $20{,}000 \times N(t)$ simulated eigenvalues. For each point in time, we can determine the biggest of the observed eigenvalues as a simple estimator of $\hat{\lambda}_{\max}(t)$ for our finite correlation estimation windows.

Equipped with finite-sample estimators for $\hat{\lambda}_{\max}$, we perform exactly the same sorting exercise as before, that is, at the end of each month, we sort the stocks into tercile portfolios according to their idiosyncratic risk and their implied skewness values and investigate the significance of the portfolio differential returns. We find no systematic deviation from our previous findings. Thus, our conclusions are essentially unchanged and robust to potential finite-sample biases.

6 Conclusion

Our study contributes to a recent strand in the literature that investigates the relationship between idiosyncratic risk and future stock returns. Previous studies typically rely on historical time series to identify systematic and idiosyncratic risk. Using a sample of all S&P 100 constituents, we identify positive risk premia for option-implied idiosyncratic risk. Depending on the model used to identify unsystematic risk, we observe a statistical and economically significant average annual premium of 1.72 percent. As the definition of unsystematic risk is closely tied to the definition of the pricing kernel, our results may be driven by misspecified asset pricing models. Therefore, we control for higher moments of the implied return distribution, specifically by implied skewness. Using a double-sorting procedure, we show that the compensation of unsystematic risk is at least partly driven by firms with high positive implied skewness.

References

Ang, A., Hodrick, R., Xing, Y., Zhang, X.: The cross-section of volatility and expected returns. J. Finance **61**(1), 259–299 (2006)

Ang, A., Hodrick, R., Xing, Y., Zhang, X.: High idiosyncratic volatility and low returns: international and further US evidence. J. Financ. Econ. **91**(1), 1–23 (2009)

Baik, J., Ben Arous, G., Péché, S.: Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. Ann. Probab. **33**(5), 1643–1697 (2005)

Bakshi, G., Madan, D.: A theory of volatility spreads. Manag. Sci. **52**(12), 1945–1956 (2006)

Bakshi, G., Kapadia, N., Madan, D.: Stock return characteristics, skew laws, and the differential pricing of individual equity options. Rev. Financ. Stud. **16**(1), 101–143 (2003)

Barberis, N., Huang, M., Santos, T.: Prospect theory and asset prices. Q. J. Econ. **116**(1), 1–53 (2001)

Beckers, S.: The constant elasticity of variance model and its implications for option pricing. J. Finance **35**(3), 661–673 (1980)

Blair, B., Poon, S.-H., Taylor, S.: Forecasting S&P 100 volatility: the incremental information content of implied volatilities and high-frequency index returns. J. Econom. **105**(1), 5–26 (2001)

Bollerslev, T., Tauchen, G., Zhou, H.: Expected stock returns and variance risk premia. Rev. Financ. Stud. **22**(11), 4463–4492 (2009)

Burda, Z., Jarosz, A., Nowak, M., Jurkiewicz, J., Papp, G.: Applying free random variables to random matrix analysis of financial data. Part I: the gaussian case. Quant. Finance **11**(7), 1103–1124 (2011)

Buss, A., Vilkov, G.: Option-implied correlation and factor betas revisited. Working Paper (2011)

Canina, L., Figlewski, S.: The informational content of implied volatility. Rev. Financ. Stud. **6**(3), 659–681 (1993)

Carhart, M.: On the persistence of mutual fund performance. J. Finance **52**(1), 57–82 (1997)
Day, T., Lewis, C.: Stock market volatility and the information content of stock index options. J. Econom. **52**(1–2), 267–287 (1992)
DeMiguel, V., Garlappi, L., Uppal, R.: Optimal versus naive diversification: how efficient is the 1/N portfolio strategy? Rev. Financ. Stud. **22**(5), 1915–1953 (2009)
DeMiguel, V., Plyakha, Y., Uppal, R., Vilkov, G.: Improving portfolio selection using Option-Implied volatility and skewness. Working Paper (2011)
Diavatopoulos, D., Doran, J., Peterson, D.: The information content in implied idiosyncratic volatility and the Cross-Section of stock returns: evidence from the option markets. J. Futures Mark. **28**(11), 1013–1039 (2008)
Driessen, J., Maenhout, P., Vilkov, G.: The price of correlation risk: evidence from equity options. J. Finance **64**(3), 1377–1406 (2009)
Drożdż, S., Kwapień, J., Grümmer, F., Ruf, F., Speth, J.: Quantifying the dynamics of financial correlations. Physica A **299**(1–2), 144–153 (2001)
Fama, E., French, K.: Common risk factors in the returns on stocks and bonds. J. Financ. Econ. **33**(1), 3–56 (1993)
Fama, E., MacBeth, J.: Risk, return, and equilibrium: empirical tests. J. Polit. Econ. **81**(3), 607–636 (1973)
Focardi, S., Fabozzi, F.: The Mathematics of Financial Modeling & Investment Management. Wiley, New York (2009)
Fu, F.: Idiosyncratic risk and the cross-section of expected stock returns. J. Financ. Econ. **91**(1), 24–37 (2009)
Geman, S.: A limit theorem for the norm of random matrices. Ann. Probab. **8**(2), 252–261 (1980)
Goetzmann, W., Kumar, A.: Equity portfolio diversification. Rev. Finance **12**(3), 433–463 (2008)
Harvey, C., Siddique, A.: Conditional skewness in asset pricing tests. J. Finance **55**(3), 1263–1295 (2000)
Jiang, G., Tian, Y.: The model-free implied volatility and its information content. Rev. Financ. Stud. **18**(4), 1305–1342 (2005)
Kollo, T.O., Ruul, K.: Approximations to the distribution of the sample correlation matrix. J. Multivar. Anal. **85**(2), 318–334 (2003)
Kraus, A., Litzenberger, R.: Skewness preference and the valuation of risk assets. J. Finance **31**(4), 1085–1100 (1976)
Krishnan, C., Petkova, R., Ritchken, P.: Correlation risk. J. Empir. Finance **16**(3), 353–367 (2009)
Laloux, L., Cizeau, P., Bouchard, J.-P., Potters, M.: Noise dressing of financial correlation matrices. Phys. Rev. Lett. **83**(7), 1467–1470 (1999)
Laloux, L., Cizeau, P., Potters, M., Bouchard, J.-P.: Random matrix theory and financial correlations. Int. J. Theor. Appl. Finance **3**(3), 391–397 (2000)
Lamoureux, C., Lastrapes, W.: Forecasting stock-return variance: toward an understanding of stochastic implied volatilities. Rev. Financ. Stud. **6**(2), 293–326 (1993)
Latané, H., Rendleman, R.: Standard deviations of stock price ratios implied in option prices. J. Finance **31**(2), 369–381 (1976)
Ledoit, O., Wolf, M.: Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. J. Empir. Finance **10**(5), 603–621 (2003)
Ledoit, O., Wolf, M.: A well-conditioned estimator for large-dimensional covariance matrices. J. Multivar. Anal. **88**(2), 365–411 (2004)
Lehar, A., Scheicher, M., Strobl, G.: Trade versus time series based volatility forecasts: evidence from the austrian stock market. Financ. Mark. Portf. Manag. **15**(4), 500–515 (2001)
Lehmann, B.: Residual risk revisited. J. Econom. **45**(1–2), 71–97 (1990)
Lintner, J.: The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. Rev. Econ. Stat. **47**(1), 13–37 (1965)
Malevergne, Y., Sornette, D.: Collective origin of the coexistence of apparent random matrix theory noise and of factors in large sample correlation matrices. Physica A **331**(3–4), 660–668 (2004)
Marčenko, V., Pastur, L.: Distribution of eigenvalues for some sets of random matrices. Math. USSR Sb. **1**(4), 457–483 (1967)
Merton, R.: A simple model of capital market equilibrium with incomplete information. J. Finance **42**(3), 483–510 (1987)
Meucci, A.: Managing diversification. Risk **May**, 74–79 (2009)
Plerou, V., Gopikrishnan, P., Rosenow, B., Nunes Amaral, L., Guhr, T., Stanley, H.: Random matrix approach to cross correlations in financial data. Phys. Rev. E **65**(6), 066126 (2002)
Pojarliev, M., Polasek, W.: Portfolio construction by volatility forecasts: does the covariance structure matter? Financ. Mark. Portf. Manag. **17**(1), 103–116 (2003)

Rajkovic, M.: Extracting meaningful information from financial data. Physica A **287**(3–4), 383–395 (2000)
Schmalensee, R., Trippi, R.: Common stock volatility expectations implied by option premia. J. Finance **33**(1), 129–147 (1978)
Sharifi, S., Crane, M., Shamaie, A., Ruskin, H.: Random matrix theory for portfolio optimization: a stability approach. Physica A **335**(3–4), 629–643 (2004)
Vanden, J.: Option coskewness and capital asset pricing. Rev. Financ. Stud. **19**(4), 1279–1320 (2006)
Zhu, J.: Pricing volatility of stock returns with volatile and persistent components. Financ. Mark. Portf. Manag. **23**(3), 243–269 (2009)

Stephan Süss Stephan Süss obtained his Doctoral degree in Finance at the University of St. Gallen. Since 2009, he has been assistant professor at the Swiss Institute of Banking and Finance. His main research interest includes risk management and the pricing of derivative products, as well as computational finance issues.